Stat 524
10-19-17
1

Multicollinearity

What if the columns of X are linearly dependent?

Then $\quad X\vec{z} = \vec{0} \quad$ for some $\vec{z} \neq \vec{0}$

But then $\quad X'X\vec{z} = X'\vec{0} = \vec{0} \quad \therefore$ Columns of $X'X$ are lin dep.

Conversely, if $X'X\vec{z}=0$ then $\vec{z}\,'X'X\vec{z} = 0$

$$\vec{z}\,'X'X\vec{z} = (X\vec{z})'(X\vec{z}) = \|X\vec{z}\|^2 = 0 \Rightarrow X\vec{z} = \vec{0}$$

$\therefore$ cols of X are lin dep.

2

That is, X is less than full rank

$\iff X'X$ is less than full rank

Detection:

1. Compute all pairwise correlations between your predictors

   In SAS, PROC CORR

2. Compute Variance Inflation Factors (VIFs)

   The VIF for the i^{th} predictor is $\dfrac{1}{1-R_i^2}$, ③

where R_i^2 is the coefficient of determination for the regression of X_i on all of the other predictors

If a VIF > 10, then that predictor is problematic

Procedure: Run full model
Remove the worst offender if its VIF > 10
Re-run the model
& repeat until all remaining VIFs ≤ 10

④

$$X = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}_{n \times p}$$

$$= \Big[\underbrace{\begin{matrix} 1 & x_{11} \cdots x_{1c} \\ \vdots & \vdots \quad\quad \vdots \\ 1 & x_{n1} \cdots x_{nc} \end{matrix}}_{\substack{X_1 \\ n \times (c+1)}} \Big| \underbrace{\begin{matrix} x_{1(c+1)} & \cdots & x_{1k} \\ \vdots & & \vdots \\ x_{n(c+1)} & \cdots & x_{nk} \end{matrix}}_{\substack{X_2 \\ n \times (k-c)}} \Big]$$

5

$$X = [X_1 | X_2]$$

$$X'X = [X_1|X_2]'[X_1|X_2]$$

$$= \left[\begin{array}{c} X_1' \\ \hline X_2' \end{array}\right][X_1|X_2]$$

$$= \left[\begin{array}{c|c} X_1'X_1 & X_1'X_2 \\ \hline X_2'X_1 & X_2'X_2 \end{array}\right]$$

Suppose that every column of X_1 is $\perp$ to every column of X_2

6

Then $X'X = \left[\begin{array}{c|c} X_1'X_1 & 0 \\ \hline 0 & X_2'X_2 \end{array}\right]$

and $(X'X)^{-1} = \left[\begin{array}{c|c} (X_1'X_1)^{-1} & 0 \\ \hline 0 & (X_2'X_2)^{-1} \end{array}\right]$

So $\hat{\beta} = (X'X)^{-1}X'Y = \left[\begin{array}{c|c} (X_1'X_1)^{-1} & 0 \\ \hline 0 & (X_2'X_2)^{-1} \end{array}\right]\left[\begin{array}{c} X_1' \\ \hline X_2' \end{array}\right]Y$

$$= \left[\begin{array}{c|c} (X_1'X_1)^{-1} & 0 \\ \hline 0 & (X_2'X_2)^{-1} \end{array}\right]\left[\begin{array}{c} X_1'Y \\ \hline X_2'Y \end{array}\right] = \left[\begin{array}{c} (X_1'X_1)^{-1}X_1'Y \\ \hline (X_2'X_2)^{-1}X_2'Y \end{array}\right]$$

7

Implication: If the columns of X_1 are $\perp$ to the columns of X_2, then deleting all of X_2 from the model will leave the parameter estimates for X_1 unchanged

---

<u>Residuals</u> Recall our model:

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

$$\text{or } Y = X\beta + \vec{\varepsilon}$$

$$E(\vec{\varepsilon}) = \vec{0}, \quad Cov(\vec{\varepsilon}) = \sigma^2 I = \begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix}$$

8

The i^{th} observed residual is

$$e_i = y_i - \hat{y}_i$$

$$\vec{e} = \begin{bmatrix} y_1 - \hat{y}_1 \\ \vdots \\ y_n - \hat{y}_n \end{bmatrix} = Y - \hat{Y} = Y - HY$$
$$= (I - H)Y$$

$$E[\vec{e}] = E[(I-H)Y] = (I-H)E[Y]$$
$$= (I-H)E[X\beta + \vec{\varepsilon}] = (I-H)[X\beta + \vec{0}]$$

$$= X\beta - HX\beta = X\beta - \underbrace{X(X'X)^{-1}X'}_{H}X\beta$$

$$= X\beta - X\beta = \vec{0}$$

$$\begin{aligned}
Cov[\vec{e}] &= Cov[(I-H)\,Y] \\
&= (I-H)\,Cov(Y)\,(I-H)' \\
&= (I-H)\,Cov[X\beta + \vec{\varepsilon}\,]\,(I-H) \\
&= (I-H)\,Cov[\vec{\varepsilon}\,]\,(I-H) \\
&= (I-H)\,\sigma^2 I\,(I-H) = \sigma^2 (I-H)^2 \\
&= \sigma^2 (I-H)
\end{aligned}$$

$\therefore\ Var(e_i) = \sigma^2(1 - h_{ii})$

and $Cov(e_i, e_j) = -\sigma^2 h_{ij}$

<u>Definitions</u>: $d_i = \dfrac{e_i}{\sqrt{MSE}}$ = standardized residual

$r_i = \dfrac{e_i}{\sqrt{MSE(1-h_{ii})}}$ = studentized residual

(11)

$$e_{(i)} = y_i - \hat{y}_{(i)} = \text{PRESS residuals}$$

predicted y-value,
but with i^{th} data value
suppressed

PRESS = "prediction error sum of squares"

Fact: $e_{(i)} = \dfrac{e_i}{1-h_{ii}}$

What if we standardize the PRESS residual? (12)

$$V[e_{(i)}] = \left(\frac{1}{1-h_{ii}}\right)^2 V[e_i]$$

$$= \left(\frac{1}{1-h_{ii}}\right)^2 \sigma^2 (1-h_{ii}) = \frac{\sigma^2}{1-h_{ii}}$$

To standardize $e_{(i)}$,
$$\frac{e_{(i)}}{\sqrt{\hat{V}[e_{(i)}]}} = \frac{\dfrac{e_i}{1-h_{ii}}}{\sqrt{\dfrac{MSE}{1-h_{ii}}}}$$

$$= \frac{e_i}{\sqrt{MSE(1-h_{ii})}} = \text{studentized residual}$$

Stat 564 HW#4

Problems 3.8, 3.9

(These are completely garbled in my PDF copy of the book, so I can’t post them here)